

Sparse Reconstruction Automaton for Synthetic Aperture Radar Tomography

Nan Ge ⁽¹⁾, Xiao Xiang Zhu ^(1,2)

(1) Remote Sensing Technology Institute (IMF)
German Aerospace Center (DLR)
Oberpfaffenhofen, 82234 Wessling, Germany
Email: {xiao.zhu; nan.ge}@dlr.de

(2) Helmholtz Young Investigators Group “SiPEO”
Chair of Remote Sensing Technology
Technische Universität München
Arcisstraße 21, 80333 Munich, Germany

Abstract—In this paper, we report our findings on automating the sparse reconstruction process in tomography with synthetic aperture radar. Two hyperparameter-free approaches are introduced into the framework of SLIMMER (Scale-down by L1 norm Minimization, Model selection, and Estimation Reconstruction). By means of numerical simulations, we evaluate their performance regarding mean and standard deviation of elevation estimates, as well as detection rate. Preliminary results with real data are also provided.

Keywords—SAR tomography; sparse reconstruction; super-resolution; SLIMMER; LASSO; SPICE; LIKES

I. INTRODUCTION

Modern spaceborne synthetic aperture radar (SAR) sensors, e.g., TerraSAR-X (TSX), TanDEM-X (TDX) and COSMO-SkyMed, acquire images with up to meter- or even decimeter-level resolution. Due to their side-looking nature, echoes from objects at the same distance are mapped onto one azimuth-range pixel, which is also known as the layover effect. On this account, interpretation of SAR images w.r.t. urban areas, where buildings are often superimposed on ground or lower infrastructures, is not straightforward. In order to separate different potential scatterers within one resolution cell and to reconstruct true three-dimensional images, tomographic SAR (TomoSAR) comes to play.

The TomoSAR technique usually employs a stack of images acquired in a repeat-pass fashion, which are characterized by slightly different sensor positions. Let us denote the axis normal to the azimuth-range imaging plane as elevation s . We work with the following well-established model which can be found in, e.g., [1]

$$\mathbf{g} = \mathbf{R}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \quad (1)$$

where $\mathbf{g} \in \mathbb{C}^{N \times 1}$ is the observation vector irregularly sampled at N orbital positions with each having a perpendicular baseline b_n w.r.t. master scene, $\mathbf{R} \in \mathbb{C}^{N \times L}$ is the sensing matrix with $R_{nl} = \exp(-j2\pi\xi_n s_l)$, $\boldsymbol{\gamma} \in \mathbb{C}^{L \times 1}$ is the reflectivity vector at L discrete elevation positions over an elevation extent Δs , with each $\gamma_l = \gamma(s_l)$ corresponding to the elevation position s_l , and $\boldsymbol{\varepsilon} \in \mathbb{C}^{N \times 1}$ is additive measurement noise, usually modeled as a complex circular Gaussian zero-mean random process. Furthermore, $\xi_n = 2b_n/\lambda r_0$ denotes spatial (elevation)

frequency, where λ and r_0 are radar wavelength and master slant range distance, respectively. Note that we have neglected other terms accounting for motion and the like without loss of generality.

Normally, we have $N \ll L$ which renders (1) an underdetermined system. To estimate $\boldsymbol{\gamma}$, additional assumptions must be made. Fortunately, it is reasonable to expect in urban scenarios that inside each resolution unit of very high resolution SAR images, only a small number of, say K , phase centers are available within Δs [2]. E.g., in case façade and ground interact with each other, $K = 2$. In other words, $\boldsymbol{\gamma}$ has only K non-zero entries and hence is K -sparse.

Similar to azimuth resolution, the elevation counterpart ρ_s is inversely proportional to elevation aperture size (baseline extent) Δb , i.e., $\rho_s = \lambda r_0 / (2\Delta b)$. For TSX/TDX, ρ_s is around 30 m and thus much worse than the meter-level azimuth and range equivalent if high resolution spotlight mode is concerned. Therefore, super-resolution power is of great interest, particularly for urban areas with drastic height changes.

In the next section, we will introduce different methodologies to pursue this goal.

II. METHODOLOGIES

A. The SLIMMER Algorithm

Scale-down by L1 norm Minimization, Model selection, and Estimation Reconstruction (SLIMMER) has been proposed in [3], [4] as a generic algorithm for sparse signal reconstruction. It consists mainly of the following three steps:

1) Scale-down by L_1 norm minimization

To exploit the sparse prior, the Least Absolute Shrinkage and Selection Operator (LASSO) problem is solved [5]

$$\hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\gamma}} \left\{ \frac{1}{2} \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda_K \|\boldsymbol{\gamma}\|_1 \right\}, \quad (2)$$

where λ_K is a hyperparameter balancing model goodness of fit and sparsity of $\boldsymbol{\gamma}$. LASSO is known to excel at identifying the support $\text{supp } \boldsymbol{\gamma} := \{1 \leq l \leq L : \gamma_l \neq 0\}$. By means of $\text{supp } \hat{\boldsymbol{\gamma}}$, the dimension of the original problem can be downscaled by extracting only the corresponding columns of \mathbf{R} . However, $\hat{\boldsymbol{\gamma}}$

can contain outliers which leads to $|\text{supp } \hat{\gamma}| > K$. Besides, (2) is prone to biased amplitude estimate. Considering these issues, the next two steps are deemed necessary.

2) Model selection

As far as the “right” model is concerned, i.e., the size of support, the model goodness of fit must be punished by its complexity to avoid overfitting. This is essentially equivalent to the following optimization problem

$$\hat{K} = \arg \min_K \left\{ -2 \ln p(\mathbf{g} | \hat{\boldsymbol{\theta}}(K), K) + 2C(K) \right\}, \quad (3)$$

where $p(\mathbf{g} | \hat{\boldsymbol{\theta}}(K), K)$ is the likelihood function of \mathbf{g} given the estimate of the unknown $\boldsymbol{\theta}(K)$ obtained with K , and $C(\cdot)$ denotes the penalty function. By choosing an appropriate $C(\cdot)$, (3) is solved to estimate the most likely \hat{K} and thus the corresponding elevation positions $\hat{\mathbf{s}}$, which may further shrink \mathbf{R} .

3) Parameter estimation

Based on dimension reduction in the first two steps, (1) is transformed into

$$\mathbf{g} = \mathbf{R}(\hat{\mathbf{s}}) \boldsymbol{\gamma}(\hat{\mathbf{s}}) + \mathbf{e}, \quad (4)$$

where \mathbf{e} contains both additive noise and model error. In case a sufficient number of observations are provided, i.e., $N \geq \hat{K}$, (4) can be solved with L_2 norm minimization

$$\hat{\boldsymbol{\gamma}}(\hat{\mathbf{s}}) = \arg \min_{\boldsymbol{\gamma}(\hat{\mathbf{s}})} \left\{ \|\mathbf{g} - \mathbf{R}(\hat{\mathbf{s}}) \boldsymbol{\gamma}(\hat{\mathbf{s}})\|_2^2 \right\}, \quad (5)$$

which concludes the SL1MMER algorithm.

The Cramér-Rao lower bound (CRLB) of elevation estimate has been derived for $K = 1, 2$ as $\sigma_{s_q,0}$ and σ_{s_q} , respectively [4][6]. Given a signal-to-noise ratio (SNR) of 3 dB, $\sigma_{s_q,0}$ is about 1 m for TSX/TDX, while σ_{s_q} varies as a function of elevation distance between two scatterers. The performance of SL1MMER algorithm has been shown to approach this lower bound [4].

Solving the LASSO problem in (2) involves predefining λ_K , which is not an easy task. In fact, λ_K can only be explicitly determined when the problem has already been solved. In real-world applications, this is done empirically, e.g., by letting

$$\hat{\lambda}_K = \sigma \sqrt{2 \ln(L)}, \quad (6)$$

where σ is the standard deviation of \mathbf{g} [5]. However, estimating σ can be as challenging as choosing λ_K , in particular with small N .

In the next subsection, we will briefly revisit two methods which are free of hyperparameter and thus require no tuning per se.

B. SPICE and LIKES

Sparse Iterative Covariance-based Estimation (SPICE), and LIKelihood-based Estimation of Sparse parameters (LIKES) have been proposed in [7].

Let $\mathbf{A} := [\mathbf{R}, \mathbf{I}_N]$ and $\mathbf{x} := [\boldsymbol{\gamma}^T, \boldsymbol{\varepsilon}^T]^T$, (1) can be rewritten as $\mathbf{g} = \mathbf{A}\mathbf{x}$. Under the assumption that the elements of \mathbf{x} are i.i.d. complex circular Gaussian zero-mean random variables, its covariance matrix can be defined as $\mathbf{C}_{\mathbf{x}\mathbf{x}} := \text{diag}(\mathbf{p})$ where $\mathbf{p} = [p_1, \dots, p_{L+N}]^T$ and p_k is the variance of x_k . By means of error propagation, the covariance matrix of \mathbf{g} is given as $\mathbf{C}_{\mathbf{g}\mathbf{g}} = \mathbf{A}\mathbf{C}_{\mathbf{x}\mathbf{x}}\mathbf{A}^H$ and its maximum-likelihood estimate is $\hat{\mathbf{C}}_{\mathbf{g}\mathbf{g}} = \mathbf{g}\mathbf{g}^H$.

On one hand, SPICE solves the following convex problem via covariance matrix fitting

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \left\{ \left\| \mathbf{C}_{\mathbf{g}\mathbf{g}}^{-1/2}(\mathbf{p}) (\mathbf{C}_{\mathbf{g}\mathbf{g}}(\mathbf{p}) - \hat{\mathbf{C}}_{\mathbf{g}\mathbf{g}}) \right\|_F^2 \right\}. \quad (7)$$

Given \mathbf{p} , \mathbf{x} can be derived to be $\mathbf{x} = \mathbf{C}_{\mathbf{x}\mathbf{x}}\mathbf{A}^H\mathbf{C}_{\mathbf{g}\mathbf{g}}^{-1}\mathbf{g}$ [7]. Eq. (7) is equivalent to

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \left\{ \mathbf{g}^H \mathbf{C}_{\mathbf{g}\mathbf{g}}^{-1}(\mathbf{p}) \mathbf{g} + \sum_{k=1}^{L+N} w_k^2 p_k \right\}, \quad (8)$$

where $w_k = \|\mathbf{a}_k\|_2 / \|\mathbf{g}\|_2$ with the square of it being the weight for \mathbf{p} , and \mathbf{a}_k is the k^{th} column of \mathbf{A} .

On the other hand, LIKES minimizes the negative log-likelihood function

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \left\{ \mathbf{g}^H \mathbf{C}_{\mathbf{g}\mathbf{g}}^{-1}(\mathbf{p}) \mathbf{g} + \ln |\mathbf{C}_{\mathbf{g}\mathbf{g}}(\mathbf{p})| \right\}, \quad (9)$$

which is a non-convex problem due to the fact that $\ln |\mathbf{C}_{\mathbf{g}\mathbf{g}}|$ is concave. By replacing $\ln |\mathbf{C}_{\mathbf{g}\mathbf{g}}|$ with its tangential plane, (9) can be iteratively approached by solving a series of SPICE-like problems with adaptive weights

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \left\{ \mathbf{g}^H \mathbf{C}_{\mathbf{g}\mathbf{g}}^{-1}(\mathbf{p}) \mathbf{g} + \sum_{k=1}^{L+N} \tilde{w}_k^2 p_k \right\}, \quad (10)$$

where $\tilde{w}_k = \mathbf{a}_k^H \tilde{\mathbf{C}}_{\mathbf{g}\mathbf{g}}^{-1} \mathbf{a}_k$, and $\tilde{\mathbf{C}}_{\mathbf{g}\mathbf{g}}$ is built with the most recent $\hat{\mathbf{p}}$.

Although no sparsity is explicitly enforced, (8) and (10) still promote sparse solutions as a result of the weighted L_1 norm of \mathbf{p} , which is known for being sparsity-driven. We integrate SPICE and LIKES into SL1MMER via simply replacing (2) by (7) or (9).

III. EXPERIMENTS USING SIMULATED AND REAL DATA

First of all, we evaluate the performance of the aforementioned methods using numerical simulations. For this purpose, we synthesize façade-ground interactions ($K = 2$) with elevation distance within ρ_s . Basically, two scenarios are taken into account with $N = 21$ which is a typical case for TomoSAR; SNR = 10, 3 dB which are considered as an upper and a lower

bound for persistent scatterers, respectively [8]. For each interaction 1000 independent realizations are generated.

As a reference, we assume that the true noise standard deviation is already known (thus avoiding any estimation error of it) and solve (2) with the λ_K defined in (6). We will refer to this as LASSO+ σ in the following. Note that in reality the true value of σ is almost never known. In this context, we merely use LASSO+ σ to establish an upper bound for any LASSO problem using (6).

Fig. 1 shows the normalized true elevation $\alpha := s/\rho_s$ of simulated façade and ground w.r.t. to their distance $\delta\alpha$ as solid lines. In addition, true elevation \pm CRLB is drawn as dashed lines for the case $N = 21$ and SNR = 10 dB.

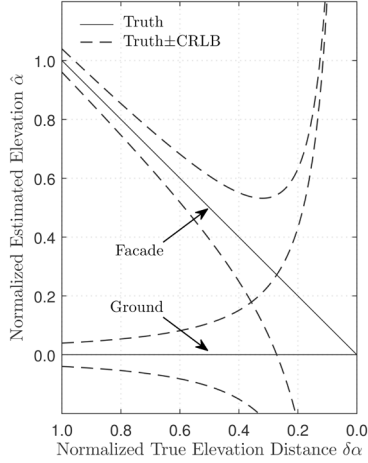


Fig. 1. True elevation of façade and ground (solid lines), and their true elevation \pm CRLB (dashed lines) with $N = 21$ and SNR = 10 dB, both w.r.t. their elevation distance.

Fig. 2 and Fig. 3 demonstrate elevation estimates with different methods and two noise levels, respectively. Each dot depicts mean elevation estimate and is accompanied with the corresponding standard deviation as error bar. Missing dots suggest that detection rate is below 20%, c.f. Fig. 4. Clearly,

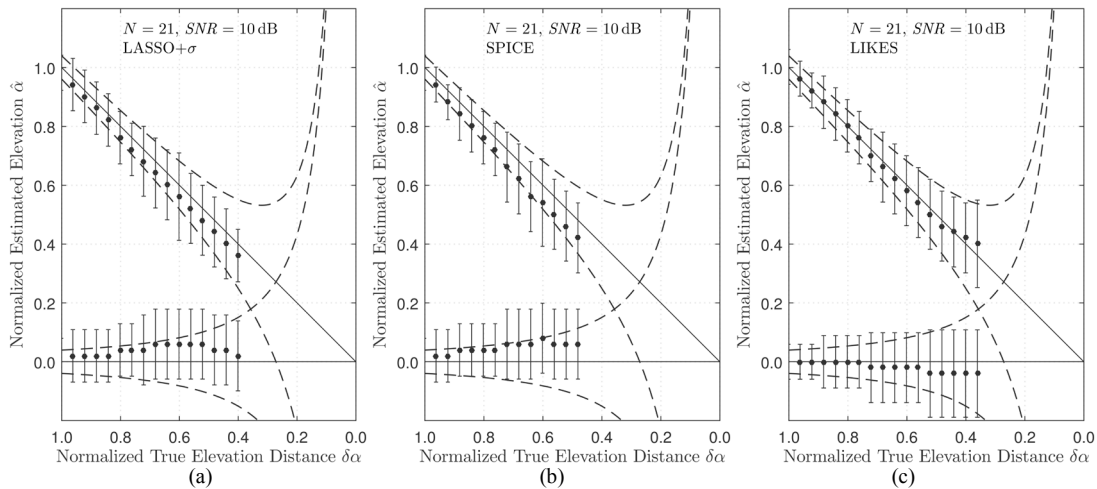


Fig. 2. Mean (dot) and standard deviation (bar) of elevation estimates for simulated façade and ground with $N = 21$ and SNR = 10 dB, (a) LASSO+ σ , (b) SPICE, and (c) LIKES.

LIKES outperforms the other two methods. The performance of LASSO+ σ and SPICE is quite similar.

We apply LIKES to a stack of 21 real bistatic SAR interferograms. Fig. 5 shows our preliminary results—the first (b) and second layer (c) of elevation estimates, superimposed on mean intensity image (a). The layover effect of roof and façade is obvious to observe at the top of the façade.

ACKNOWLEDGMENT

This work is supported by the the Helmholtz Association under the framework of the Young Investigators Group “SiPEO” (VH-NG-1018, www.sipeco.bgu.tum.de). The authors would like to thank Steffen Peter for valuable discussions.

REFERENCES

- [1] G. Fornaro, F. Serafino, and F. Soldovieri, “Three-dimensional focusing with multipass SAR data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 507–517, Mar. 2003.
- [2] X. X. Zhu and R. Bamler, “Very High Resolution Spaceborne SAR Tomography in Urban Environment,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.
- [3] X. Zhu and R. Bamler, “Tomographic SAR Inversion by L1-Norm Regularization -- The Compressive Sensing Approach,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, 2010.
- [4] X. X. Zhu and R. Bamler, “Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation With Application to Spaceborne Tomographic SAR,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247–258, 2012.
- [5] S. Chen, D. Donoho, and M. Saunders, “Atomic Decomposition by Basis Pursuit,” *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, Jan. 1998.
- [6] R. Bamler, M. Eineder, N. Adam, X. Zhu, and S. Gernhardt, “Interferometric Potential of High Resolution Spaceborne SAR,” *Photogramm. - Fernerkund. - Geoinformation*, vol. 2009, no. 5, pp. 407–419, Nov. 2009.
- [7] P. Stoica and P. Babu, “SPICE and LIKES: Two hyperparameter-free methods for sparse-parameter estimation,” *Signal Process.*, vol. 92, no. 7, pp. 1580–1590, Jul. 2012.
- [8] N. Adam, R. Bamler, M. Eineder, and B. Kampes, “Parametric Estimation and Model Selection Based on Amplitude-Only Data in PS-Interferometry,” presented at the Fringe 2005 Workshop, 2006, vol. 610, p. 21.

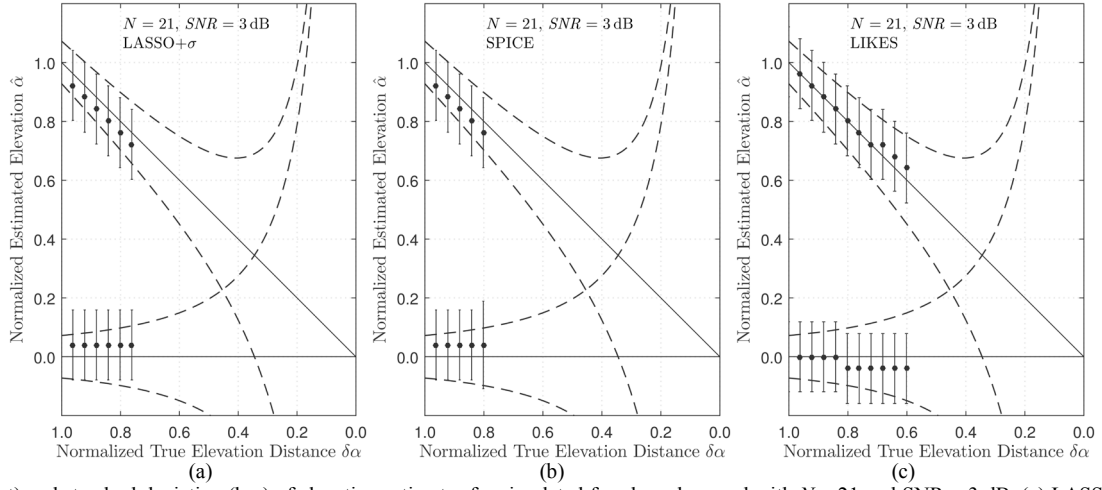


Fig. 3. Mean (dot) and standard deviation (bar) of elevation estimates for simulated façade and ground with $N = 21$ and $\text{SNR} = 3$ dB, (a) LASSO+ σ , (b) SPICE, and (c) LIKES.

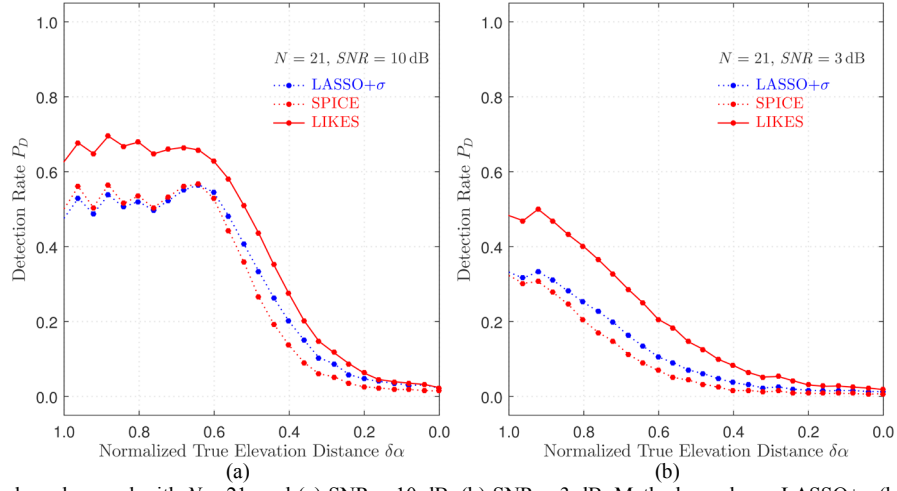


Fig. 4. Detection rate of façade and ground with $N = 21$, and (a) $\text{SNR} = 10$ dB, (b) $\text{SNR} = 3$ dB. Methods used are: LASSO+ σ (blue dotted lines), SPICE (red dotted lines), and LIKES (red solid lines).

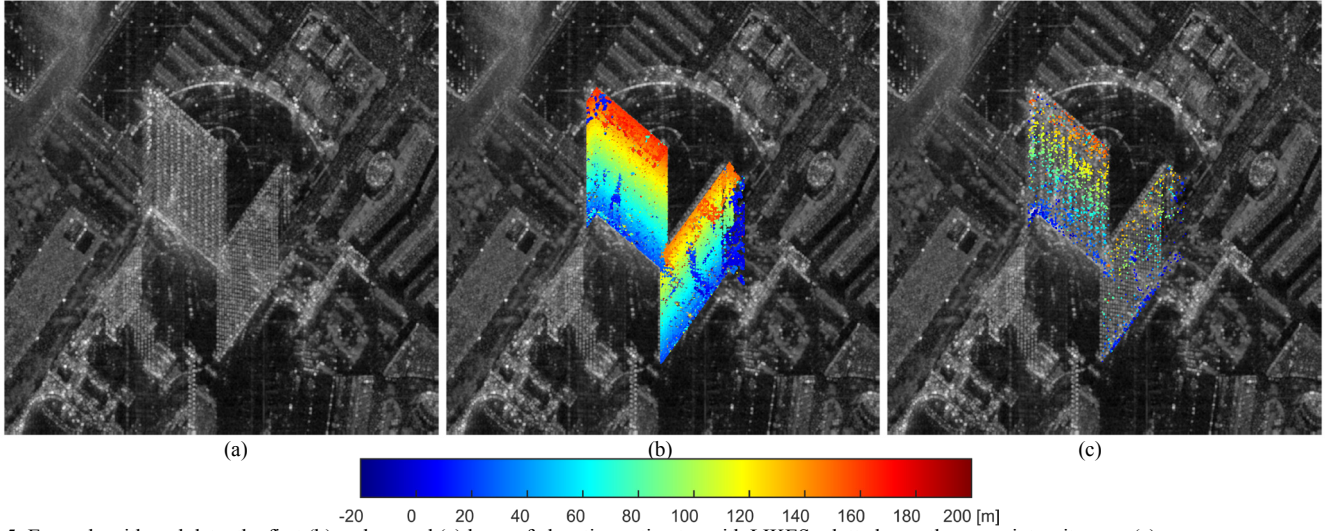


Fig. 5. Example with real data: the first (b) and second (c) layer of elevation estimates with LIKES, plotted over the mean intensity map (a).